An extended Pacific–North American index from upper-air historical data back to 1922

Ewen, Tracy ; Brönnimann, Stefan ; Annis, Jeffrey

Abstract: This paper presents a reconstruction of a Pacific–North America (PNA) index from historical upper-level data for the period 1922–47. The data have been compiled from a number of sources and cover the Pacific–North American sector relatively well over this time period. Temperature and geopotential height profiles from aircraft, kite, and radiosonde ascents back to 1922 have been digitized and validated. Wind speed and direction from pilot balloon data back to the early 1920s, provided by NCAR, have also been used. A statistical regression approach is used for the reconstruction and calibrated in the post-1948 period using NCEP–NCAR reanalysis data. Split-sample validation experiments were performed within the NCEP–NCAR period, and sensitivity experiments with different subsets of predictors were performed. Similar reconstructions and validation experiments were carried out using a 540-yr control run from the Community Climate System Model, version 3 (CCSM3). The reconstructed index series together with validation statistics for both the historical and model data are presented. Excellent reconstruction skill is found for the winter months, while the reconstructions are somewhat worse in summer. Compared with a reconstruction based only on surface data, the addition of the newly digitized upper-air stations improves the reconstruction skill in all seasons. The historical reconstruction is presented with respect to its imprint on hemispheric fields of surface air temperature, sea level pressure, and precipitation with a special focus on extreme cases. In addition, the extended PNA index is compared with indices of the North Atlantic Oscillation, the Pacific decadal oscillation, and the El Niño–Southern Oscillation. The relationship to these indices is found to be stationary over the analysis period.

DOI: https://doi.org/10.1175/2007JCLI1951.1

An Extended Pacific–North American Index from Upper-Air Historical Data Back to 1922

TRACY EWEN, STEFAN BRÖNNIMANN, AND JEFFREY ANNIS

Institute for Atmospheric and Climate Science, ETH Zurich, Zürich, Switzerland

(Manuscript received 14 March 2007, in final form 13 July 2007)

ABSTRACT

This paper presents a reconstruction of a Pacific–North America (PNA) index from historical upper-level data for the period 1922–47. The data have been compiled from a number of sources and cover the Pacific–North American sector relatively well over this time period. Temperature and geopotential height profiles from aircraft, kite, and radiosonde ascents back to 1922 have been digitized and validated. Wind speed and direction from pilot balloon data back to the early 1920s, provided by NCAR, have also been used. A statistical regression approach is used for the reconstruction and calibrated in the post-1948 period using NCEP–NCAR reanalysis data. Split-sample validation experiments were performed within the NCEP–NCAR period, and sensitivity experiments with different subsets of predictors were performed. Similar reconstructions and validation experiments were carried out using a 540-yr control run from the Community Climate System Model, version 3 (CCSM3). The reconstructed index series together with validation statistics for both the historical and model data are presented. Excellent reconstruction skill is found for the winter months, while the reconstructions are somewhat worse in summer. Compared with a reconstruction based only on surface data, the addition of the newly digitized upper-air stations improves the reconstruction skill in all seasons. The historical reconstruction is presented with respect to its imprint on hemispheric fields of surface air temperature, sea level pressure, and precipitation with a special focus on extreme cases. In addition, the extended PNA index is compared with indices of the North Atlantic Oscillation, the Pacific decadal oscillation, and the El Niño–Southern Oscillation. The relationship to these indices is found to be stationary over the analysis period.

1. Introduction

An important gap in our understanding of present climate variability and assessing future climate change inevitably remains the extension of the past climate record. In particular, the extension of the upper-air network of observations prior to the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (1948), is essential to improve our understanding of tropospheric circulation changes over the last century and the role that these changes may have had on past extreme climate events. Past droughts over the Great Plains and Canadian Prairies and their possible relation to the tropical oceans can only be understood in the context of the midtropospheric circulation over the Pacific–North American (PNA) sector (see, e.g., Namias 1978; Trenberth et al. 1988; Trenberth and Branstator 1992; Shabbar et al. 1990; Maybank et al. 1995; Trenberth and Guillemot 1996; Chen and Newman 1998; Bonsal et al. 1999; Garnett 2001). The same holds for climate extremes in the Arctic.

The PNA pattern represents one of the most prominent large-scale patterns of atmospheric low-frequency variability in the northern extratropics (e.g., Simmons 1982; Barnston and Livezey 1987). It represents a wave train spanning the North Pacific–North American region, with two north–south-oriented centers of action over the North Pacific and two northwest–southeast-oriented centers over the United States (Wallace and Gutzler 1981; Barnston and Livezey 1987). The PNA affects climate variability over North America during most of the year. Temperature variations in northwestern North America and the southwestern United States are strongly related to the PNA in winter and spring and to a lesser extent in autumn. Correlations with precipitation tend to be weaker but nonetheless show co-

Corresponding author address: Tracy Ewen, Institute for Atmospheric and Climate Science, ETH Zurich, Universitätsstrasse 16, CH-8092 Zürich, Switzerland.
E-mail: tracy.ewen@env.ethz.ch

DOI: 10.1175/2007JCLI1951.1

herent regions of high correlation throughout North America (see, e.g., Leathers et al. 1991; Robertson and Ghil 1999). The PNA has been shown to have little influence on summer variations in both surface temperature and precipitation (Leathers et al. 1991).

The corresponding PNA index is a useful measure of large-scale variations in circulation on time scales of months to decades. The index, however, is defined based on 500-hPa geopotential height (GPH) data (Wallace and Gutzler 1981; Barnston and Livezey 1987), which are currently only available back to 1948. Although the sea level pressure (SLP) fields can be used to assess a PNA-like pattern of variability further back in time (Trenberth 1990; Wallace and Thompson 2002; Quadrelli and Wallace 2004; Overland and Wang 2005), and indices such as the North Pacific index (NPI) diagnosed from SLP can be used as a proxy to assess the PNA (Deser et al. 2004; Trenberth and Hurrell 1994), data from higher atmospheric levels are required to fully understand the dynamical processes that govern large-scale patterns of variability in the middle troposphere and the corresponding downstream effects on climate over North America. The PNA pattern is shown in Fig. 1 as the point correlation between the 500-hPa height anomalies and the PNA index over the 1948–2000 period.

In this paper, we reconstruct a PNA index from reevaluated upper-air observations from 1922 to 1947. In addition to conventional cross-validation techniques, we also validate our reconstructions using model data from the NCAR Community Climate System Model, version 3 (CCSM3). The remainder of this paper is organized as follows: In section 2 the data, reconstruction, and validation methods as well as sensitivity experiments are outlined, including an overview of the reconstruction method and regression model. The reconstructed indices and the validation statistics for both historical and model data are presented in section 3, followed by an analysis of the surface climate patterns and the temporal variability in section 4. In section 5 extremely low and high PNA indices from both observations and model data are examined, and surface fields of SLP, temperature, and precipitation are compared. A summary and conclusions follow in section 6.

2. Data and methods

a. Historical upper-air and additional data

Monthly historical upper-level data that have been recently reevaluated are used for the reconstruction. Data from kites (1922–34, temperature on geometric altitudes up to around 4 km), aircraft (1925–42, temperature and later pressure on geometric altitudes up to around 5 km), and radiosondes (1937–47, temperature and geopotential height on standard pressure levels) for North America were digitized from the tables published in the *Monthly Weather Review* (Brönnimann 2003; Ewen et al. 2008). In addition, reevaluated upper-level wind data (u and v wind components on geometric altitude levels) from selected stations from the global TD52 and TD53 pilot balloon datasets for the period 1919–47 provided by NCAR (available online at http://dss.ucar.edu/docs/papers-scanned/papers.html, documents RJ0167, RJ0168) are used.

In total, 110 upper-air stations were used, which cover the Pacific–North American sector quite well (Fig. 1). Each station record covers only part of the period and Fig. 1 (bottom) shows the total number of available predictor values per month during the historical period (1922–1947). The upper air data availability is best during WWII, however some upper-air kite data are available back to 1922 and the pilot balloon data cover the earlier prewar period quite well (e.g., upper-level winds from Hawaii are available back to 1922).

The upper-level data during the historical period (1922–47) were supplemented with surface air temperature records from the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS) database (Hansen et al. 1999). In addition, we used gridded SLP data from the Hadley Centre sea level pressure dataset (HadSLP2; Allan and Ansell 2006), where the latitude-weighted principal component (PC) time series of the monthly anomaly fields between 20°N:80°N and 160°E:60°W are used. We retained 15 PCs, corresponding to approximately 96% of the total variance. The PNA index, used as the predictand in the linear regression model, is constructed using the point-wise index of the 500-hPa height fields, standardized, defined by Wallace and Gutzler (1981).

b. Reconstruction method

To reconstruct the PNA index from historical station data, statistical models need to be calibrated in a period for which both the PNA index and the station (predictor) data are available. We used the 1948–2003 period, where the PNA index can be calculated from NCEP–NCAR reanalysis data (Kalnay et al. 1996; Kistler et al. 2001). In many cases, upper-level station data are not available for the same locations over the entire 1922–2003 period, or they have long gaps. Therefore, we use NCEP–NCAR reanalysis data, interpolated to the station location, to supplement all historical upper-level data after 1948 (for the altitude level data, the reanalysis data were interpolated to geometric altitudes).

It is important to keep in mind that reanalysis data are a model product, which will differ from observa-

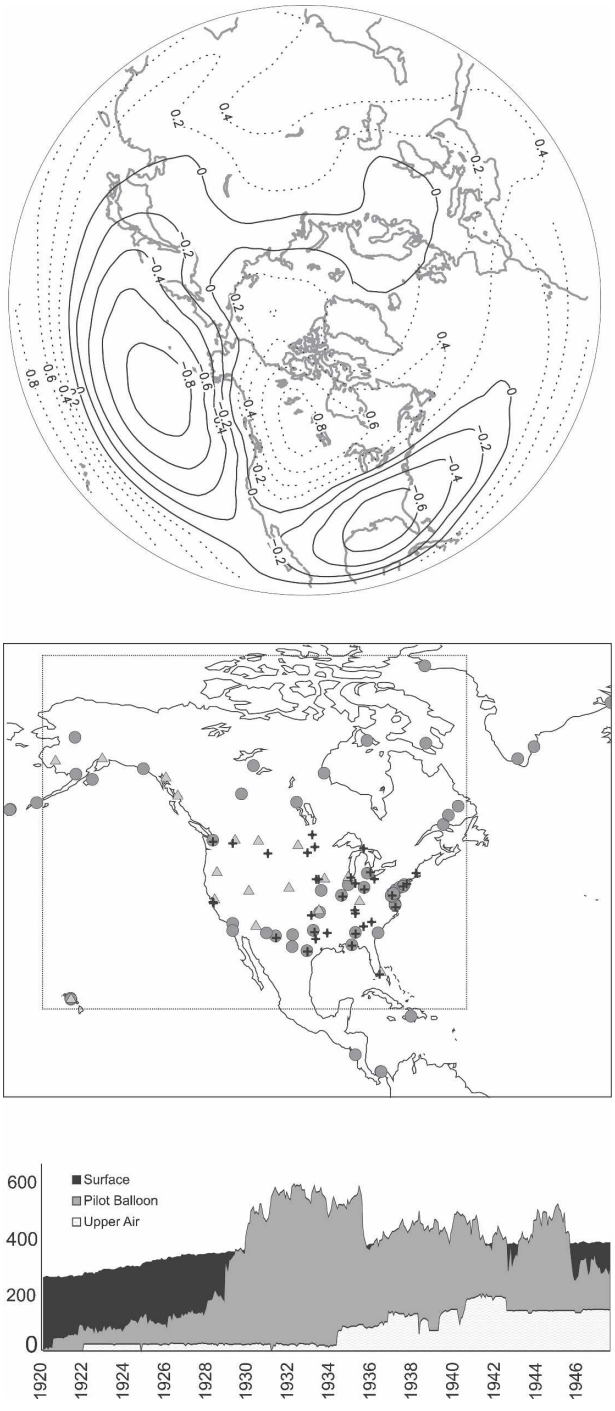


FIG. 1. (top) The PNA pattern shown as the point correlation between the 500-hPa height anomalies and the PNA index over the NCEP–NCAR reanalysis period. The correlation contour interval is 0.2. The solid line indicates negative correlations and the dashed line positive correlations. (middle) Upper-air stations used in the historical reconstruction: pilot balloon (circles), kite (triangles), and aircraft or radiosonde (crosses). The dashed line shows the area used to calculate the SLP PCs. (bottom) Total number of available predictor variables for each month over the historical reconstruction period for surface temperature (black), upper air (temperature and pressure, light gray), and pilot balloons (light gray).

tions. Additional errors arise because, for example, the quality of the historical data is lower than more recent observations and from the interpolation procedure. In an attempt to account for these types of errors, the supplemented station data (i.e., the interpolated NCEP–NCAR reanalysis data) have been randomly perturbed by normally distributed noise. The standard deviation of the noise was estimated based on our quality assessment (Brönnimann 2003; Ewen et al. 2008) in which NCEP–NCAR reanalysis was also used as a reference.

Gaps in the NASA GISS surface station data are generally short. In the calibration period they are filled with anomaly data from neighboring stations in order to have a complete data series. The PC time series from the SLP field is continuous. This results in a calibration dataset that has no missing values over the 1948–2003 period. All data series are expressed as standardized anomalies from the 1961 to 1990 mean annual cycle.

The reconstruction method is a simple principal component regression model and is based on Brönnimann et al. (2006). All historical data (upper air, pilot balloon, and surface variables) are used as predictor variables in the model. The relationship between the predictor variables and the PNA index (predictand) is then determined for the calibration period, when both the PNA index and station data are available. Due to missing values prior to the calibration period, each month in the historical period has a different combination of available variables. The reconstructions were therefore performed independently for each month from 1922 to 1947. A three-calendar month moving window was used for calibration; to form a reconstruction model for January 1922, for example, we used all Decembers, Januaries, and Februaries in the calibration period. For each time step all available variables were weighted, with 50% of the weight attributed to the surface and the remainder to the upper-level data. Within the surface data, half of the weight was attributed to the SLP PCs and half to the temperature data. Within the SLP PCs, the weight was split according to their correlation with the PNA index. Within the upper-level data, half of the weight was attributed to the pilot balloon data and the remaining half of the weight to the upper-level temperature or geopotential height data. Note that over the oceans, only SLP data are used. No additional SST data were used in the procedure.

After the weighting, a principal component analysis was performed on all available predictor variables in the calibration period to reduce the number of variables. The amount of variance retained was varied between 70% and 98% and the optimum fraction [according to the mean reduction of error (RE) from the vali-

dation experiments, see below] was chosen for each time step (i.e., each month).

The reconstructions were validated using two split-sample validation experiments, that is, the calibration period was split and one part was used for calibrating the model and the other for validating the model. In the first experiment we used 1948–66 for validation and in the second 1985–2003. The validation statistic used is the RE (Cook et al. 1994) averaged from both validation experiments. The reduction of error is defined as

$$\mathrm{RE} = 1 - \frac{\Sigma_t (x_{\mathrm{rec}} - x_{\mathrm{obs}})^2}{\Sigma_t (x_{\mathrm{obs}})^2},$$

where t is time, x_{obs} is the observed value, and x_{rec} is the reconstructed value. Values of RE can be between $-\infty$ and 1 (perfect reconstructions). For RE values greater than around 0.2, RE is close to R^2, the coefficient of determination. Here, RE > 0, determined in an independent period, is normally considered an indication that the model has predictive skill.

Sensitivity experiments were performed using subsets of the data: only data from the earth's surface, only upper-air data, or only pilot balloon data. For brevity, these experiments will be referred to as reconstruction with surface data (recon_surf), reconstruction with all upper-air data, including pilot balloons, (recon_ua), reconstruction with pilot balloon data (recon_pball), and reconstruction with surface, upper-air, and pilot balloon data (recon_all).

c. CCSM3

Split-sample validations are a logical way to test properties of the statistical models used for the reconstructions. They are, however, based on shorter calibration periods and hence may not provide an accurate measure of the reconstruction skill. An alternative is to use climate model data and simulate the reconstruction procedure within these data (e.g., von Storch et al. 2004). For this purpose, we used data from the NCAR CCSM3 [Collins et al. (2006); downloaded from the Earth System Grid Web site http:/www.earthsystemgrid.org]. We used the 540-yr control run b30.009 performed at a resolution of T85 (approximately 1.4° × 1.4°), which includes prescribed (i.e., fixed to 1990 levels) concentrations of greenhouse gases, aerosol species (sulfate, dust, carbon, sea salt, and volcanic), fixed solar irradiance, and ozone.

The major modes of variability are well represented in CCSM3. El Niño–Southern Oscillation (ENSO) variability has been diagnosed using the Niño-3.4 index and considerable tropical variability is found; the amplitude and zonal extent of variability in the equatorial Pacific SSTs is well simulated (Deser et al. 2006; Collins et al. 2006). The North Atlantic Oscillation (NAO) pattern and index are also realistically represented (Holland 2003; Kiehl and Gent 2004), and the spatial pattern of the 500-hPa height anomalies is well simulated; however, the anomalies are weaker by a factor of 2 (Deser et al. 2006). Precipitation in the southeastern United States is underestimated in the CCSM3 model and slightly overestimated in the Pacific Northwest when compared to the Global Precipitation Climatology Project (GPCP) dataset (Xie et al. 2003; Collins et al. 2006). Surface temperature is generally well represented in the model; however, there are large biases in the 2-m air temperature in the sub-Arctic and in Alaska when compared to observations (Collins et al. 2006).

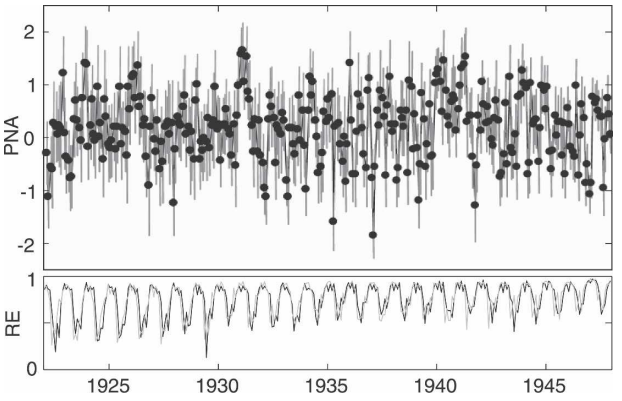


FIG. 2. (top) The reconstructed monthly PNA index for the 1922–47 period with 95% confidence intervals calculated from the two split-sample validation experiments. (bottom) Reduction of error (RE) for the two split-sample validation experiments.

For the reconstruction, monthly means of temperature, geopotential height, and u, v winds were extracted from the postprocessed data, by interpolating to the exact location of the historical upper-air stations. The data were degraded by adding noise as described for the NCEP–NCAR reanalysis data. The reconstructions, validation, and sensitivity experiments were then simulated in five subperiods over the 540-yr model run.

3. Results

a. Historical reconstructions

Figure 2 shows the reconstructed PNA index for the 1922–47 period with 95% confidence intervals, calculated from the two split-sample validation experiments, together with the RE statistics for each month. The reconstruction captures variability over both the monthly and interannual time scales. The amplitude

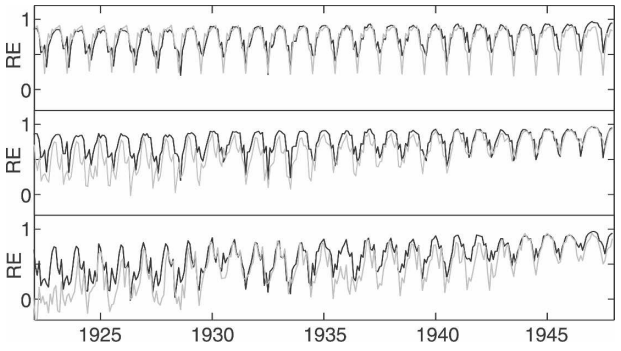


FIG. 3. (top) Reconstruction RE values for all data (recon_all, black) and surface stations only (recon_surf, gray). (middle) RE values for all data (recon_all, black) and upper-air data (recon_ua, gray) and (bottom) for all upper-air data (recon_ua, black) and pilot balloon data only (recon_pball, gray).

and variance of the historical reconstruction is similar to the NCEP–NCAR period (variance of 0.42 versus 0.47). There is one extremely low value in January 1937, which is robust with respect to the amount of variance retained (two even lower values were observed after 1948). High indices are found in the winter and spring of 1930–31 over a 4-month period, in winter and spring of 1925–26, and for both winters of 1939–40 and 1940–41. These large positive anomalies appear to be robust features of the reconstruction. In fact, all four winters coincide with strong El Niño events. The negative values in the autumns of 1926 and 1941, however, also coincide with El Niño events. The mean PNA index over the entire reconstruction period is 0.18.

The RE statistics show a clear annual cycle. The reconstructions are extremely good in late winter (with RE values ranging between 0.90 and 0.96), whereas they are less reliable in summer (0.35 to 0.60). In addition to monthly changes, the RE values change over time due to the different number of predictor variables used. During World War II, a large amount of upper-air data is available and the reconstructions are quite good even in the summer months. Overall, the RE values obtained from the two split-sample validations (black, gray) are similar.

To understand whether there is a benefit to adding the historical upper-air data, RE values from the four reconstructions are compared (recon_all, recon_surf, recon_ua, and recon_pball). Figure 3 (top) shows that the full reconstruction (recon_all, black curve) and the surface-only reconstruction (recon_surf, gray curve) are very similar over the first part of the period. After around 1928 the reconstruction improves (in terms of RE) when including upper-air data, particularly in summer, and remains slightly better than the surface reconstruction over the period, even in winter. In Fig. 3 (middle) the RE values for recon_all (black curve) are compared to recon_ua (gray curve). Until around 1941, surface temperature and SLP data are still needed (in addition to the upper-air and pilot balloon data) to obtain the best possible reconstructions over all months. After that time, the radiosonde network was sufficiently dense, and upper-air-only reconstructions perform better in summer months, although for the other months, both reconstructions are quite similar. Comparing the RE values in Fig. 3 (bottom) for recon_ua and recon_pball reveals that the inclusion of the upper-air temperature and pressure data considerably improves the reconstruction skill.

Altogether, the upper-air data are needed as the elevated topography in the region of the anomaly center located over the cordillera of northwestern North America tends to destroy the coherency of both the geopotential height and pressure fields at lower elevations (Blackmon et al. 1979; Leathers et al. 1991). Our reconstructions suggest that this is particularly the case in summer. Although the PNA is not a major mode of variability during the summer months, it is still a rough measure of the stationary waves over North America and the inclusion of upper-air data contributes information about the general character of the flow.

b. Validation with CCSM3 model data

Before performing validation experiments in the CCSM3 control run, the model calculated PNA should be analyzed. Although 500-hPa GPH anomalies are too weak (see above), this does not necessarily affect the PNA, which we have defined based on standardized anomalies of 500-hPa GPH at four grid points. In fact, the variance of the PNA index in the CCSM3 control run is 0.59, which is larger than in the reanalysis period (0.47). Also, extremes are slightly larger in the model. The imprint of the PNA on the surface climate variables (see also sections 4 and 5) is in good agreement with NCEP–NCAR reanalysis. Hence, we think that the control run climate is adequate to serve as a substitute in which to test our reconstruction approach.

To validate our reconstruction approach in the model, the 540-yr control run was broken into five adjacent 108-yr periods, each replicating the period 1900–2007. Within each subperiod, the same reconstruction and validation procedures were applied as for the reconstructions of the real atmosphere, including flagging those values as missing where no data for the real atmosphere are available in the corresponding month and adding random noise to the predictors. The result (Fig. 4) is a collection of five independent PNA reconstructions ("model-reconstructed index"), each of which can

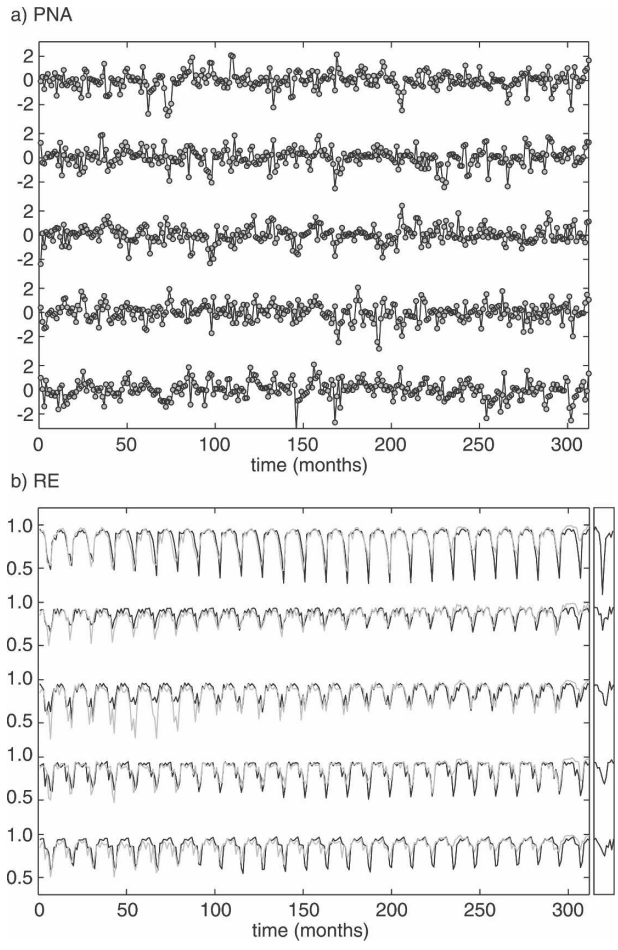


FIG. 4. (a) Five reconstructed intervals of the PNA using CCSM3 model output. (b) Mean RE values for each interval from split-sample validations. Both the full reconstruction (recon_all, gray curve) and upper-air reconstruction (recon_ua, black curve) are shown. The side plot shows the corresponding RE values calculated per calendar month for the entire reconstruction periods.

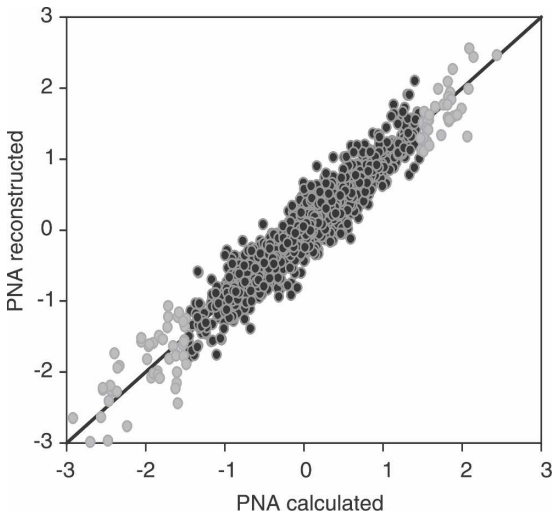


FIG. 5. Model-reconstructed series vs model-calculated series for all five reconstruction periods, $R^2 = 0.92$. Extreme indices (defined as ±2 standard deviations of the mean, $R^2 = 0.97$) are shown in light gray revealing fewer positive than negative extremes.

be compared with the corresponding "true" PNA index calculated from the model 500-hPa fields ("model-calculated index") over the model reconstruction period. The statistics can be compared to the "historical-reconstructed index" (recon_all) and the "reanalysis-calculated index."

The amplitudes of the five model-reconstructed PNA time series (Fig. 4a) show large anomalies between ±2.0–3.0, slightly larger than the largest anomalies found in either the historical-reconstructed or the reanalysis-calculated indices, which are around ±2.5. The variances for the model-reconstructed indices are similar over each interval (around 0.55), slightly smaller than for the model-calculated indices (0.59). Figure 5 shows the model-reconstructed versus model-calculated indices. The explained variance is 0.92, which is relatively high. Extreme indices are shown in light gray and are defined as values outside ±2 standard deviations from the mean. There are more negative than positive extremes, which can also be seen in the time series. The extremes seem to be well reproduced in the model ($R^2 = 0.97$).

We also simulated the split-sample validation procedure. Figure 4b shows the mean RE values (average of both split-sample validations) for each period using only upper-air data (black curves) and using both surface and upper-air data (gray curves). The RE values are all quite good and range from around 0.3 to 0.7 in the summer months to consistently high values between 0.90 and 0.95 during the winter months, which progressively improve as more upper-air station data become available over time. These RE values are very similar to those found in the split-sample validation from the reanalysis period. Interestingly, in the model split sample validations the upper-air-only reconstructions are often better than those with both surface and upper-air data, whereas this is only the case for summers after 1941 in the reanalysis split-sample validations.

On the right side of Fig. 4b we added the RE values obtained from comparing the model-reconstructed values with the model-calculated values calculated over each of the five reconstruction periods and averaged by calendar month. Comparing this with the left side of the figure shows the difference between the uncertainty estimated from split-sample validations and the true uncertainty. The agreement is relatively good for the absolute values of RE in all five periods. RE is mostly

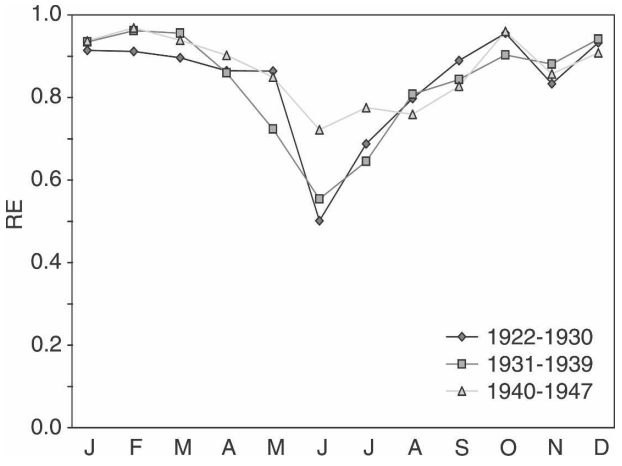


FIG. 6. RE values using different predictor data (kites, pilot balloons, aircraft, and radiosondes) calculated from the five reconstruction periods in the CCSM3 validation experiments averaged per calendar month.

between 0.8 and 0.97 except in summer, where values are as low as 0.2–0.6. Interestingly, the seasonal cycles of the RE values from the split-sample validations capture the seasonal variation of the true RE relatively well. This shows that split-sample validations within the calibration period give a good impression of the uncertainty in the reconstructions. By implication, this means that the split-sample validations performed within the NCEP–NCAR data (Fig. 3) are a good measure for the uncertainty of the historical reconstructions.

Although the RE values change over each period, as the networks of the predictor data are prescribed to match the historical reconstruction, this is not considered in the RE values shown in Fig. 4b (right). In Fig. 6 we show RE values calculated per calendar month over all five reconstruction periods but separated by three different stages of the development of the upper-air networks. We distinguish the period 1922–30, when the upper-air predictors are mostly pilot balloon and kite data, 1931–39 (pilot balloon and aircraft), and 1940–47 (pilot balloon and radiosonde). The differences between the network periods are small during the cold season. Clear differences can be seen in the summer season. At this time of the year, the additional levels in the radiosonde data (kites and aircraft barely reach 5 km) help to obtain better reconstructions.

4. Analysis of robustness, stationarity, and low-frequency variability

In this section we analyze the variability of the reconstructed PNA index and its relation to fields of climate variables at the earth's surface to assess the robustness of the reconstructions. Moreover, by comparing reconstructions with independent data, we address issues of stationarity and low-frequency variability.

Figures 7a–c show the correlation between surface air temperature (CRUTEM2; Jones and Moberg 2003) and the reconstructed PNA over both the reconstructed (Figs. 7a,b) and NCEP–NCAR reanalysis periods (Fig. 7c) from January to March. Similarly, correlation maps for the Global Historical Climatology Network (GHCN), version 2 precipitation dataset (Peterson and Vose 1997) and the reconstructed PNA are shown (Figs. 7d–f) over both periods.

The PNA imprint during the reanalysis period has been thoroughly investigated by others (see above). In the positive phase, the PNA pattern (Fig. 1) exhibits centers of action over western Canada (higher than normal pressure) and the southeastern United States (lower than normal pressure). Warmer temperatures on the west coast of North America and increased precipitation in the southwestern United States occur during this phase. There is a strong gradient across the middle of North America where the correlations are close to zero. Correlations close to zero do not necessarily mean that there is no imprint, but that the imprint is not always the same due to shifts in the anomaly centers. During a positive phase of the PNA index, temperatures in the Southeast tend to be below normal due to the development of a deep trough over the Southeast, which allows polar air masses to move into the South more frequently. During a negative PNA phase the temperatures tend to be above normal in this region. Correlations for precipitation tend to be weaker than those for temperature. As for temperature, the mechanisms associated with changes in precipitation during the positive phase are related to the deep trough that extends into the eastern United States. This tends to carry cyclones much farther south and east than usual, resulting in decreased precipitation in the mid-Southeast and an increase along the southeastern coast. The introduction of polar air masses into this region also inhibits the transport of moisture inland from the Gulf of Mexico. Over the Rockies, the high ridge tends to push Pacific storms northward, decreasing precipitation.

Our reconstruction approach is based on predictor variables that directly capture the deep troughs, cold-air outbreaks, and cyclones rather than the 500-hPa wave pattern. Hence, we determine the PNA via its effects on temperature and pressure. As the reconstruction method assumes stationarity, we also expect to find the same correlation patterns in the reconstruction period (1922–47) as in the reanalysis. In fact, the spatial

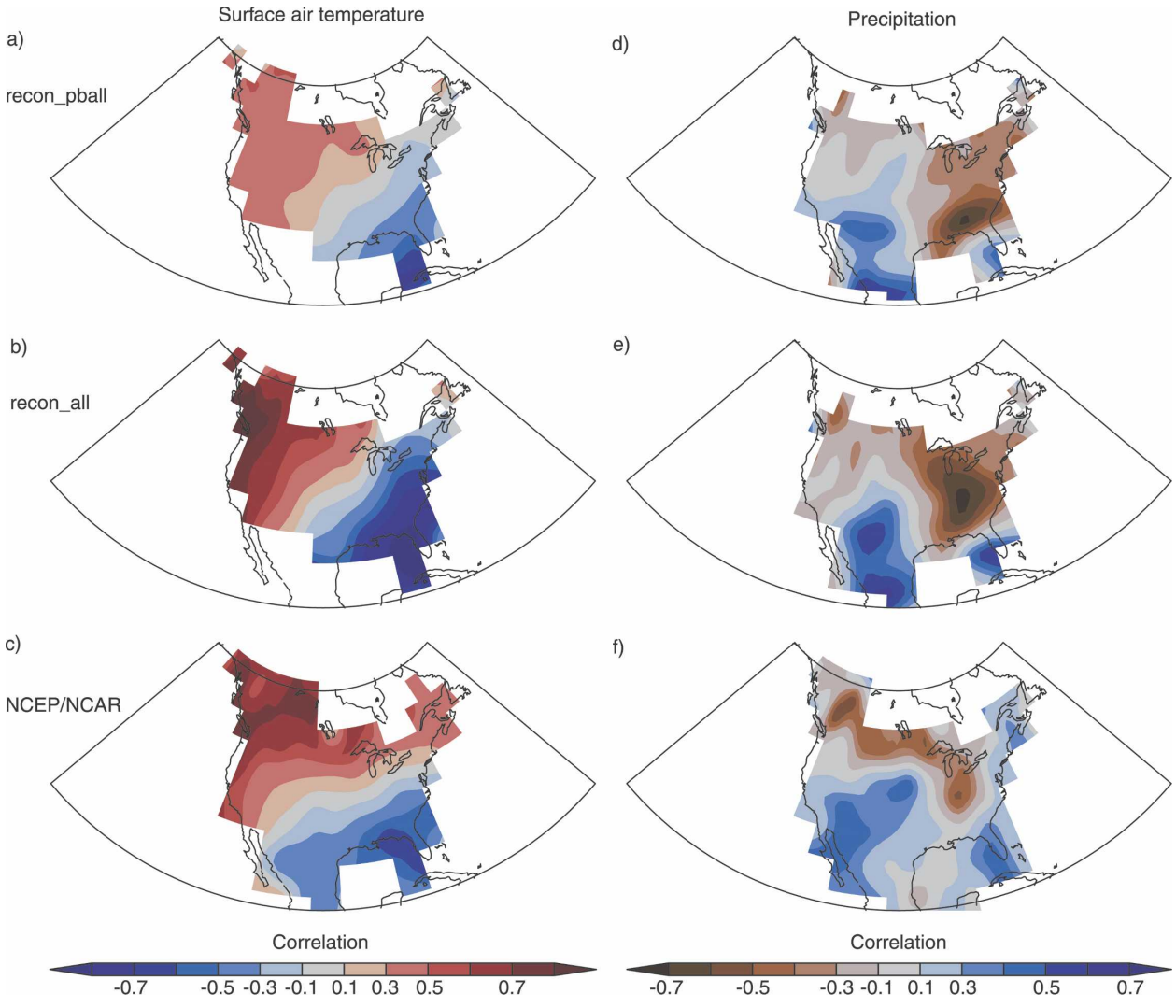


FIG. 7. Correlation maps over North America of the winter PNA (January–March) with (a)–(c) surface temperature from the CRUTEM2v dataset and (d)–(f) precipitation from the GHCN, version 2. Maps for the reconstructed period (1922–47) are shown in (a), (d) (recon_pball) and (b), (e) (recon_all) and the NCEP–NCAR reanalysis period (1948–2000) in (c), (f).

pattern of temperature is quite similar (Figs. 7b,c). The reconstructed PNA index is not directly dependent on precipitation data and we find similar spatial patterns in the correlation maps (Figs. 7e,f). The assumption of stationarity, however, could also be wrong. Several other authors have used upper-air indices over the reanalysis period to track changes in midtropospheric flow over North America and found that a major shift took place during the late 1950s (Leathers and Palecki 1992). We can evaluate the stationarity assumption to some extent (i.e., remove the direct dependence of the reconstructions on surface temperature and pressure) by using the PNA index reconstructed with only pilot balloon winds (recon_pball). Although we find slightly lower correlations (Fig. 7a) than for the full reconstruction (Fig. 7b), the spatial pattern is very similar. Hence, we have no reason to assume that the PNA effects on surface temperature could have been nonstationary.

The correlation maps for precipitation (Figs. 7e,f) do show differences between the periods. In particular, the band of negative correlations that extends west of the Great Lakes, across the northern United States and into western Canada, found in the reanalysis period (Fig. 7f), is not apparent in the reconstruction period with pilot balloons (Fig. 7d). The spatial pattern of positive correlations is also different between the periods, and the feature extending from the Southwest into the central United States in the reanalysis period (Fig. 7f) is quite different from that in the reconstruction period (Fig. 7d), which does not extend to the Southwest coast,

including the Baja peninsula. Taken together, this suggests that there may have been a shift in variability between the two periods, using only upper-air wind data.

We also analyzed the relations between the PNA index and more remote climate modes, such as ENSO, the Pacific decadal oscillation (PDO), and NAO. In this case, the stationarity of the relation can be better addressed because the assumption of local stationarity, on which the reconstruction is based, does not necessarily require a stationarity of the remote teleconnections. Arguably the most important remote teleconnection with the PNA is that related to ENSO. In Fig. 8a, a winter PNA index (November–February average) is compared to a Niño-3.4 index with a lead time of 2 months (September–December average). The relationship between ENSO and the PNA is similar in both the reconstruction and the reanalysis period. El Niño explains 38% of the winter-to-winter PNA variability in the reanalysis period and 27% in the historical period. As discussed above, the most prominent anomaly features of the reconstructed PNA index on an interannual scale are related to strong El Niño events. Apart from interannual variability, the PNA index also shows decadal variability, with pronounced minima in the late 1940s and early 1970s.

Of particular interest in this context is the relation between the PNA and the PDO, which is addressed in Fig. 8b. For this purpose, the wintertime PNA index and an annual PDO index were smoothed with a Gaussian filter ($\sigma = 3$, cutoff period of 16 yr). The figure clearly shows that the low-frequency variability of the PNA index is closely related to the PDO over both periods. Even though the nature of the PDO low-frequency variability is debated (Newman et al. 2003), this shows that the atmosphere follows the same low-frequency variability as the North Pacific SSTs.

The extended PNA index may also help to analyze Pacific–Atlantic coupling. While a "seesaw" between the Aleutian and Icelandic lows has been demonstrated (Honda et al. 2001), the relationship between the PNA and NAO is still unclear. Figure 8c shows the PNA index (November–February) together with an NAO index (January–April) that lags the PNA by 2 months. The correlation between the PNA and NAO in the reanalysis period is −0.20, which is not significant at the 95% level. In the reconstruction period, the correlation is stronger, that is, −0.29. While this alone is also not significant (due to the low number of degrees of freedom), it provides some independent confirmation for a weakly negative relationship. Over the entire period, the correlation is statistically significant.

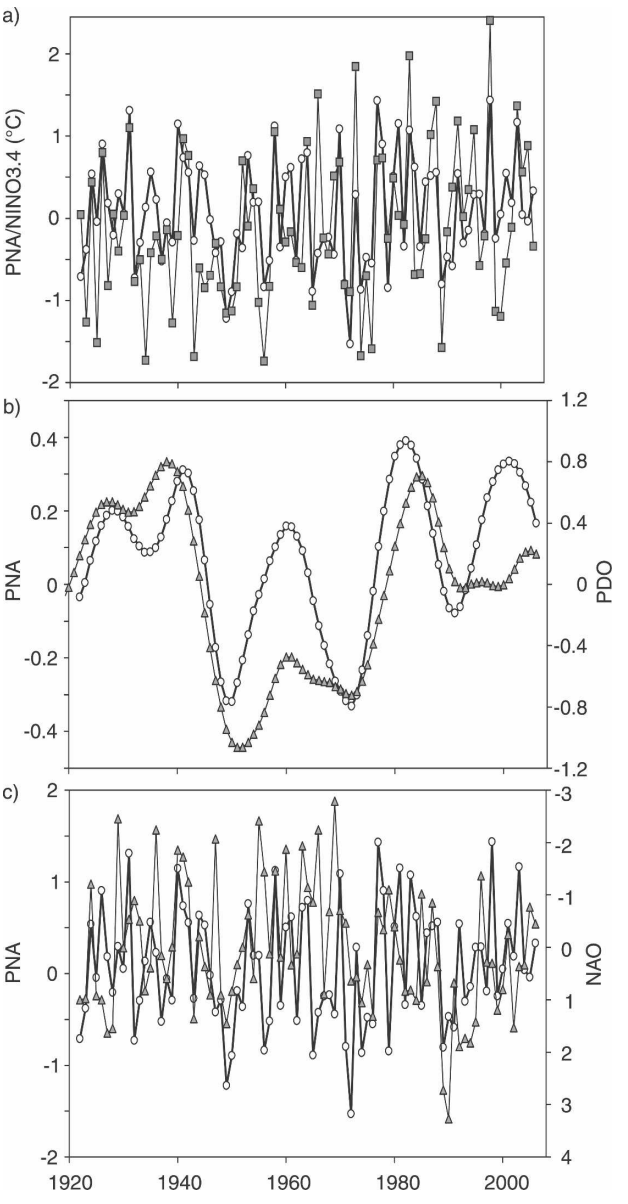


FIG. 8. A winter PNA index (November–February) is compared to (a) a Niño-3.4 index (September–December), (b) an annual PDO index, and (c) an NAO index (January–April). In (b), both the PNA and PDO are smoothed with a Gaussian filter ($\sigma = 3$). The PNA is shown as white circles in each.

5. Extreme indices

Climate extremes are of particular interest in climate research and any reconstruction of the PNA index should also adequately capture extremes. As shown above, extremes of the PNA index are well reproduced when our reconstruction approach is applied to climate model data. In this section, we investigate extremes in the PNA reconstructions in more detail. In particular, we analyze whether or not there is a possible asymme-

try (of the patterns) and nonlinearity (of the amplitudes) of the climate imprint of extreme PNA indices.

Our reconstruction approach is linear. Multiplying each predictor with a constant results in a PNA index multiplied by the same constant. Asymmetries or nonlinearities, if present, therefore could hamper the validity of our reconstructions. To address this point, we analyzed months with extreme indices (i.e., outside two standard deviations of the CCSM3 model calculated index, the NCEP calculated index, and the reconstructed PNA index, respectively). Composites of surface temperature [the Hadley Centre Climatic Research Unit Temperature version 2 (HadCRUT2v)], SLP (HadSLP2), and precipitation (GHCN) anomalies were calculated for both positive and negative extremes (Fig. 9). In the reanalysis period, composites for positive and negative extremes are almost symmetric over the Pacific–North American sector, both with respect to the patterns as well as the amplitudes. This suggests that there are no asymmetries or nonlinearities in the relation between the PNA and the climate imprint that could invalidate our reconstructions. A similar result is found in the CCSM3 model (note that here we use only the five reconstruction periods). No asymmetries are found in the Pacific–North American sector, which means that we can safely address the performance of the reconstructions in the model even for analyzing extremes. The magnitude of both the SLP and 2-m temperature anomalies, however, is larger in CCSM3 than in the NCEP–NCAR reanalysis. The SLP composite for positive extremes exhibits a negative anomaly centered in the Pacific, bounding the coast of North America. The positive anomaly to the west is centered slightly farther northeast than a typical wintertime positive phase PNA pattern, which is normally centered along the Pacific Northwest, extending up into northern Canada. The pattern, however, is quite typical of a positive phase PNA in winter as seen in the surface field (Quadrelli and Wallace 2004). The amplitude of the anomalies of the SLP field is almost 2 times larger in the CCSM3 composites than in the historical and NCEP–NCAR reanalysis periods. Note that the variability is underestimated in CCSM3 for 500-hPa GPH but not for SLP. This discrepancy may be due to the control run having fixed external forcings (Kiehl and Gent 2004), whereas in reality the climate forcings are changing over time.

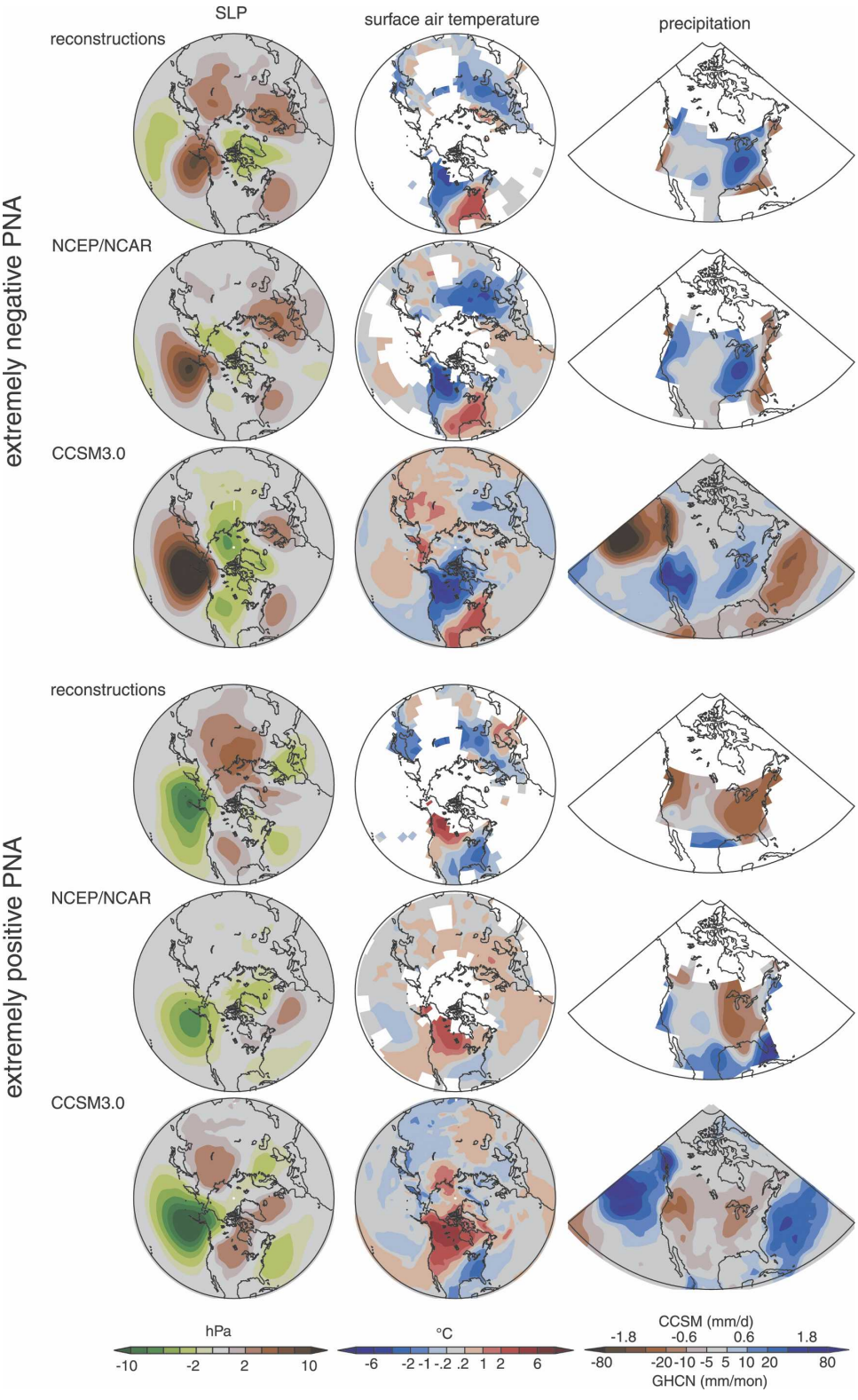
It is interesting, however, to analyze more remote regions. In the NCEP–NCAR period, we find that positive extremes project onto an NAO-like SLP pattern over the North Atlantic, while negative extremes show a strong SLP anomaly over western Europe. Surface air temperature shows corresponding nonlinearities over Eurasia. In CCSM3, we find no clear NAO pattern in the composites for positive extremes, but we find a similar anomaly center in the SLP field (in the negative case) over western Europe. Despite some differences between NCEP–NCAR reanalysis and the CCSM3 control run over Eurasia, this analysis shows that the pattern related to PNA extremes is symmetric and linear in the Pacific–North American region, but can be asymmetric in more remote regions, such as Eurasia.

We can now turn to extremes found in the historical PNA reconstructions. These extremes are robust features; they do not depend on the fraction of variance retained, weighting, or different combinations of predictor variables (surface data only versus surface and upper-air data). This indicates that for the study of extreme PNA indices, surface data should be sufficient to properly reconstruct these events.

The historical composites for positive and negative extremes (Fig. 9) also show symmetry over the Pacific–North American sector, as is to be expected from the linear reconstruction approach. The SLP field shows negative pressure anomalies (positive PNA composite) over the main centers of action; however, the positive anomalies (negative PNA composite) are slightly weaker in both the reanalysis and reconstruction than CCSM. The negative SLP anomaly over northwestern North America is not present in the reconstruction and quite weak in the reanalysis as compared to CCSM. The temperature field also shows symmetry with the typical temperature pattern over North America (negative PNA composite). The positive PNA composite, however, shows somewhat weaker anomalies than the negative PNA composite. Interestingly, the negative PNA

→

FIG. 9. Composites of SLP (HadSLP2), surface air temperature and SST anomalies (HadCRUT2v), and precipitation (GHCN) anomalies for winter months with high or low PNA indices. Composites are based on either the historical reconstructed PNA (1922–47), reanalysis calculated PNA (1948–2002), or CCSM3 model calculated PNA (for 130 yr from model run). (top three rows) Composites of negative anomalies for the historical-reconstructed index, the reanalysis-calculated index, and the model-calculated index. (bottom three rows) Positive anomalies of the same three PNA indices. Each composite is calculated with indices ±2 standard deviations of the mean of the PNA index for each interval, for winter months. The negative composites include 8 events (historical reconstructed), 15 events (reanalysis calculated), and 51 events (model calculated). The positive composites include 8 events (historical reconstructed), 11 events (reanalysis calculated), and 27 events (model calculated).


SLP
surface air temperature
precipitation
extremely negative PNA
reconstructions
NCEP/NCAR
CCSM3.0
extremely positive PNA
reconstructions
NCEP/NCAR
CCSM3.0
hPa
-10 -2 2 10
°C
-6 -2 -1 -.2 .2 1 2 6
CCSM (mm/d)
-1.8 -0.6 0.6 1.8
-80 -20 -10 -5 5 10 20 80
GHCN (mm/mon)

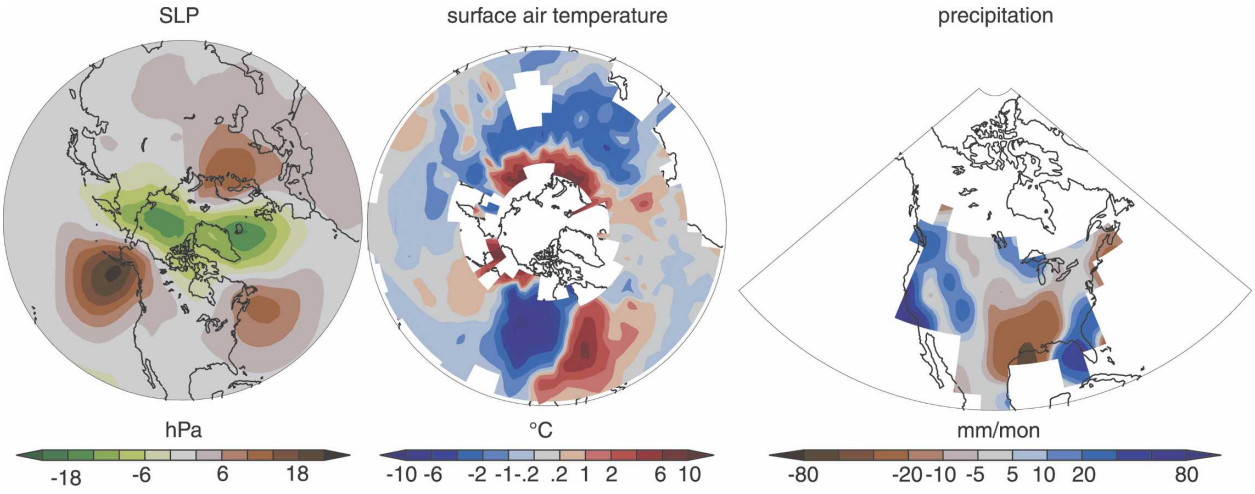


FIG. 10. HadSLP2, HadCRUT2v, and GHCN precipitation anomalies for January 1937, a month with an extremely low PNA index found in the historical reconstruction. The temperature fields are plotted on a nonlinear scale to discern ocean SSTs.

composites for all three (reconstructions, reanalysis, and CCSM3) show positive SLP anomalies and negative temperature anomalies in eastern Europe and northern Eurasia and hence are asymmetric in this respect. In the positive PNA composite, however, the temperature pattern is quite different, especially in the NCEP–NCAR reanalysis data, which shows warm anomalies over Europe and Eurasia (negative for the reconstruction and CCSM3). This might point to a somewhat different teleconnection pattern in the reconstruction period.

One interesting feature of our index is the extremely low value found in January 1937. We have analyzed Arctic winter surface temperatures using data from Polyakov et al. (2003), HadCRUT2v, and GHCN temperature and found 1937 to be one of the warmest winters in the Arctic during the warming phase from 1910 to 1945 (Polyakov et al. 2003; Overland and Wang 2005). The temperature field for this month (Fig. 10) shows much warmer than normal temperatures in the Arctic, where data are available from the HadCRUT2v dataset. This feature is much more pronounced over Siberia than in the negative composite (Fig. 9). This suggests that anomalous midtropospheric circulation may have played a role on surface climate variations over this period.

6. Summary and conclusions

In this paper, we have reconstructed a historical PNA index, prior to NCEP–NCAR reanalysis, extending back to 1922 using a newly digitized and validated upper-air dataset. This index provides new information on the state of midtropospheric circulation prior to the NCEP–NCAR reanalysis period, including changes in variability over the extended record and extreme PNA events, which may have played a role in certain climate extremes during this period.

To assess the importance of early upper-air data, we have carried out sensitivity experiments for subsets of the data. Although the surface data contribute important information in the wintertime, the reconstruction improves as upper-air data are added to the reconstruction, especially in the summer months. This is especially true as more upper-air data become available after 1940. Although the PNA is not normally considered a major mode of variability during the summer months, it is still a good measure of the stationary wave pattern over North America, which can be well characterized with the PNA reconstructions.

In addition to split-sample validations within the NCEP–NCAR period, the reconstructions were assessed by simulating the entire procedure in a 540-yr control run from the NCAR CCSM3 model. We find that the RE values, representing the reconstruction skill, are very similar to those found in the split-sample validation from the reanalysis period.

New information concerning the state of the upper-level circulation will prove valuable for studying the influence of the large-scale SST field on North American climate and the occurrence of droughts (Latif and Barnett 1994; Robertson 1996; Robertson and Ghil 1999; Hoerling and Kumar 2003; Straus and Shukla 2002; Schubert et al. 2004). It may also provide insight into 1910–45 Arctic warming (Polyakov et al. 2003; Overland and Wang 2005), which had one of the warmest winters in 1937. We find that the most extreme negative index occurs in January 1937 and is a robust

feature of our reconstruction. This extreme index corresponds to much higher than normal temperatures in the southern and southeastern United States and in the Arctic, suggesting that variations in midtropospheric flow played a role in these surface anomalies.

Another important question concerning variations in upper-level circulation is whether or not there has been a change in the variability over the observational record. We find that in general the spatial pattern of correlation maps between the extended PNA index and surface temperature (CRUTEM2v) are similar over both the historical reconstruction and the NCEP–NCAR reanalysis period. With precipitation (GHCN), however, the spatial pattern is somewhat different, suggesting that there may have been a shift in the variability between these two periods.

The relationship to other indices (Niño-3.4, NAO, and PDO) appears to be stationary over the entire analysis period. In particular, ENSO explains approximately 30% of the winter-to-winter PNA variability, and the PNA index is strongly correlated to the low-frequency signal of the PDO, indicating that the atmosphere follows the low-frequency variability of the North Pacific SSTs. The PNA and NAO show a weakly negative relationship, which is statistically significant. Further analysis is needed to fully understand these relationships and will be the focus of a following study.

Acknowledgments. This study was funded by the Swiss National Science Foundation (PP002-102731). We would like to thank Roy Jenne at NCAR for providing us with pilot balloon data from the TD52 and TD53 datasets. Additional data have been provided by the Climatic Research Unit (Norwich, United Kingdom), the Hadley Centre of the Met Office (Exeter, United Kingdom), NOAA–CIRES Climate Diagnostics Center (Boulder, Colorado), and NASA GISS (New York, New York). We also wish to thank two anonymous reviewers for helpful comments.

REFERENCES

Allan, R., and T. Ansell, 2006: A new globally complete monthly historical gridded mean sea level pressure dataset (HadSLP2): 1850–2004. *J. Climate,* **19,** 5816–5842.

Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.,* **115,** 1083–1126.

Blackmon, M. L., R. A. Madden, J. M. Wallace, and D. S. Gutzler, 1979: Geographical variations in the vertical structure of geopotential height fluctuations. *J. Atmos. Sci.,* **36,** 2450–2466.

Bonsal, B. R., X. Zhang, and W. D. Hogg, 1999: Canadian Prairie growing season precipitation variability and associated atmospheric circulation. *Climate Res.,* **11,** 191–208.

Brönnimann, S., 2003: A historical upper air-data set for the 1939–44 period. *Int. J. Climatol.,* **23,** 769–791.

——, T. Ewen, T. Griesser, and R. Jenne, 2006: Multidecadal signal of solar variability in the upper troposphere during the 20th century. *Space Sci. Rev.,* **125,** 305–317.

Chen, P., and M. Newman, 1998: Rossby wave propagation and the rapid development of upper-level anomalous anticyclones during the 1988 U.S. drought. *J. Climate,* **11,** 2491–2504.

Collins, W., and Coauthors, 2006: The Community Climate System Model Version 3 (CCSM3). *J. Climate,* **19,** 2122–2143.

Cook, E. R., K. R. Briffa, and P. D. Jones, 1994: Spatial regression methods in dendroclimatology: A review and comparison of two techniques. *Int. J. Climatol.,* **14,** 379–402.

Deser, C., A. S. Phillips, and J. W. Hurrell, 2004: Pacific interdecadal climate variability: Linkages between the Tropics and North Pacific during boreal winter since 1900. *J. Climate,* **17,** 3109–3124.

——, A. Capotondi, R. Saravanan, and A. S. Phillips, 2006: Tropical Pacific and Atlantic climate variability in CCSM3. *J. Climate,* **19,** 2451–2481.

Ewen, T., A. Grant, and S. Brönnimann, 2008: A monthly upper-air dataset for North America back to 1922 from the *Monthly Weather Review*. *Mon. Wea. Rev.,* in press.

Garnett, R., 2001: The Canadian Prairie drought of 2001: A four billion dollar shortfall. *CMOS Bull.,* **30,** 37–39.

Hansen, J., R. Ruedy, J. Glascoe, and M. Sato, 1999: GISS analysis of surface temperature change. *J. Geophys. Res.,* **104,** 30 997–31 022.

Hoerling, M., and A. Kumar, 2003: The perfect ocean for drought. *Science,* **299,** 691–694.

Holland, M., 2003: The North Atlantic Oscillation–Arctic Oscillation in the CCSM2 and its influence on Arctic climate variability. *J. Climate,* **16,** 2767–2781.

Honda, M., H. Nakamura, J. Ukita, I. Kousaka, and K. Takeuchi, 2001: Interannual seesaw between the Aleutian and Icelandic lows. Part I: Seasonal dependence and life cycle. *J. Climate,* **14,** 1029–1041.

Jones, P. D., and A. Moberg, 2003: Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *J. Climate,* **16,** 206–223.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kiehl, J. T., and P. R. Gent, 2004: The Community Climate System Model, version 2. *J. Climate,* **17,** 3666–3682.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–268.

Latif, M., and T. P. Barnett, 1994: Causes of decadel climate variability in the North Pacific/North Atlantic sector. *Science,* **266,** 634–637.

Leathers, D. J., and M. A. Palecki, 1992: The Pacific/North American teleconnection pattern and United States climate. Part II: Temporal characteristics and index specification. *J. Climate,* **5,** 707–716.

——, B. Yarnal, and M. A. Palecki, 1991: The Pacific/North American teleconnection pattern and United States climate. Part I: Regional temperature and precipitation associations. *J. Climate,* **4,** 517–528.

Maybank, J., B. R. Bonsal, K. Jones, R. Lawford, E. G. O'Brien, E. Ripley, and E. Wheaton, 1995: Drought as a disaster. *Atmos.–Ocean,* **33,** 195–222.

Namias, J., 1978: Multiple causes of the North American abnormal winter 1976–77. *Mon. Wea. Rev.,* **106,** 279–295.

Newman, M., G. P. Compo, and M. A. Alexander, 2003: ENSO-forced variability of the Pacific decadal oscillation. *J. Climate,* **16,** 3853–3857.

Overland, J. E., and M. Wang, 2005: The Arctic climate paradox: The recent decrease of the Arctic Oscillation. *Geophys. Res. Lett.,* **32,** L06701, doi:10.1029/2004GL021752.

Peterson, T. C., and R. S. Vose, 1997: An overview of the Global Historical Climatology Network temperature database. *Bull. Amer. Meteor. Soc.,* **78,** 2837–2849.

Polyakov, I. V, R. V. Bekryaev, U. S. Bhatt, R. L. Colony, A. P. Maskshtas, D. Walsh, G. V. Alekseev, and M. A. Johnson, 2003: Variability and trends of air temperature and pressure in the maritime Arctic, 1875–2000. *J. Climate,* **16,** 2067–2077.

Quadrelli, R., and J. M. Wallace, 2004: A simplified linear framework for interpreting patterns of Northern Hemisphere wintertime climate variability. *J. Climate,* **17,** 3728–3744.

Robertson, A. W., 1996: Interdecadal variability in a multicentury climate integration. *Climate Dyn.,* **12,** 227–241.

——, and M. Ghil, 1999: Large-scale weather regimes and local climate over the western United States. *J. Climate,* **12,** 1796–1813.

Schubert, S. D., M. J. Suarez, P. J. Pegion, R. D. Koster, and J. T. Bacmeister, 2004: On the cause of the 1930s Dust Bowl. *Science,* **303,** 1855–1859.

Shabbar, A., K. Higuchi, and J. L. Knox, 1990: Regional analysis of Northern Hemisphere 50 kPa geopotential heights from 1946 to 1985. *J. Climate,* **3,** 543–557.

Simmons, A. J., 1982: The forcing of planetary wave motion by tropical diabatic heating. *Quart. J. Roy. Meteor. Soc.,* **108,** 503–534.

Straus, D. M., and J. Shukla, 2002: Does ENSO force the PNA? *J. Climate,* **15,** 2340–2358.

Trenberth, K. E., 1990: Recent observed interdecadal climate changes in the Northern Hemisphere. *Bull. Amer. Meteor. Soc.,* **71,** 988–993.

——, and G. W. Branstator, 1992: Issues in establishing causes of the 1988 drought over North America. *J. Climate,* **5,** 159–172.

——, and J. W. Hurrell, 1994: Decadal atmospheric-ocean variations in the Pacific. *Climate Dyn.,* **9,** 303–319.

——, and C. J. Guillemot, 1996: Physical processes involved in the 1988 drought and 1993 flood in North America. *J. Climate,* **9,** 1288–1298.

——, G. W. Branstator, and P. A. Arkin, 1988: Origins of the 1988 North American drought. *Science,* **242,** 1640–1645.

von Storch, H., E. Zorita, J. M. Jones, Y. Dimitriev, F. Gonzalez-Rouco, and S. F. B. Tett, 2004: Reconstructing past climate from noisy data. *Science,* **306,** 679–682.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.,* **109,** 784–812.

——, and D. Thompson, 2002: The Pacific center of action of the Northern Hemisphere annular mode: Real or artifact? *J. Climate,* **15,** 1987–1991.

Xie, P., J. E. Janowiak, P. A. Arkin, R. Adler, G. J. Huffman, S. Curtis, A. Gruber, and R. Ferraro, 2003: GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *J. Climate,* **16,** 2197–2214.